Soil Biology & Biochemistry 105 (2017) 227–235

Contents lists available at ScienceDirect

Soil Biology & Biochemistry

journal homepage: www.elsevier.com/locate/soilbio

Combined use of network inference tools identifies ecologically meaningful bacterial associations in a paddy soil

Honggui Wang [a, b], Zhong Wei [c], Lijuan Mei [b], Jingxin Gu [b], Suisui Yin [b], Karoline Faust [d, e, f, 1], Jeroen Raes [d, e, f], Ye Deng [g], Yulong Wang [a], Qirong Shen [c, 1], Shixue Yin [b, *, 1]

[a] Key Laboratory of Crop Genetics and Physiology of Jiangsu Province, China
[b] College of Environmental Science and Engineering, Yangzhou University, Yangzhou 225127, China
[c] College of Resources and Environmental Science, Nanjing Agricultural University, Nanjing, 210095, China
[d] Department of Structural Biology, Flemish Institute for Biotechnology (VIB), Belgium
[e] VIB Centre for the Biology of Disease, VIB, Herestraat 49, 3000 Leuven, Belgium
[f] Laboratory of Microbiology, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussels, Belgium
[g] Institute of Research Centre for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China

ARTICLE INFO

Article history:
Received 23 September 2016
Received in revised form
26 November 2016
Accepted 30 November 2016
Available online 6 December 2016

Keywords:
Soil microbial community
Interaction network inference
Ecologically meaningful interactions
Paddy soil

ABSTRACT

High-throughput sequencing technologies have recently made it possible to interrogate the phylogenetic diversity of soils at considerable depth. This ability has led to the development of many computational tools to infer interaction networks from environmental samples. Although such tools have widely been used, they have more often served as a visual means to compare microbial communities across environmental gradients than as a means to appreciate microbial interactions associated with certain ecological processes. Previous studies have often regarded a subnetwork (module) as a functional unit but its functionality in ecological context has never been evidenced. To make better use of these tools in appreciating microbial interactions, we propose the combinational use of different inference tools. This ensemble approach permits the use of more independent predictors and the removal of tool-specific predictions in order to increase prediction accuracy. The purpose of the present study is to identify ecologically meaningful bacterial associations using multi-tool approach. Soil samples were collected in time series from experimental paddy rice plots. Bacterial communities were characterized by high-throughput tag sequencing of 16S rRNA gene fragments. We used three tools, Co-occurrence Network inference (CoNet), Molecular Ecological Network Analysis (MENA) and extended Local Similarity Analysis (eLSA), to infer networks from abundance profiles, partitioned the networks into modules, screened for the modules with $\geq$50% of genus-/species-level nodes, captured the modules that were derived from different tools and shared $\geq$ 50% of order-level nodes (tool-agreed modules) and tested their robustness against the changes in the tool parameters. By these procedures, two three-tool-agreed modules were found. One represented a guild that is phenotypically associated with aerobic respiration and fermentation and the other represented a guild phenotypically associated with metal/sulphur cycles, all of which are essential processes of water-submerged paddy soils that are mediated by bacteria. These data suggested that the linked members in a module were functionally associated taxa that work together to achieve a distinct function or an ecological process, and thus were ecologically meaningful to the environment. We selected three linked species from a three-tool-agreed module and validated their interactions using co-culture methods. Results showed that the interaction type between *Janthinobacterium lividum* and *Leuconostoc lactis* in the two-species mixture was validated to be ambivalent, positive for one partner and negative for the other. However, this type of interaction was not retained when a third party *Lactococcus piscium* was introduced, signifying the complexity of multi-species interactions. Validation results suggested that the selected species were interacting partners in laboratory but the validated interaction types were different from those inferred. By multi-tool approach, we also

* Corresponding author. Huayang West Road 196, College of Environmental Science and Engineering, Yangzhou University, Yangzhou 225127, Jaingsu Province, China.
E-mail address: sxyin@yzu.edu.cn (S. Yin).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.soilbio.2016.11.029

found that highly linked nodes, which are often referred to as "keystone species" and are frequently interpreted as the species playing important roles in soils, are tool dependent. Among top ten highly linked nodes, only four are conserved across three tools. These results suggest more research is required on the ecological significance of degree-based identification of keystone species. Overall, the present study highlights the potential utility of combined use of inference tools to identify ecologically meaningful bacterial associations in soils and other environmental samples. It is interesting to see what type of ecologically meaningful bacterial associations can be found in other soils.

1. Introduction

High-throughput sequencing technologies have recently made it possible to interrogate phylogenetic diversity in environmental samples at considerable depth of coverage. This ability has led to the development of many computational tools to infer interaction networks from complex microbial communities, on the basis of abundance profiles. Among the long list are Co-occurrence Network inference (CoNet) (Faust et al., 2012), Molecular Ecological Network Analysis (MENA) (Deng et al., 2012), extended Local Similarity Analysis (eLSA) (Xia et al., 2013), Sparse Correlations for Compositional data (SparCC) (Friedman and Alm, 2012), Learning Interactions from Microbial Time Series (LIMITS) (Fisher and Mehta, 2014), Correlation inference for Compositional data through Lasso (CCLasso) (Fang et al., 2015), Microbe-microbe Interaction Networks (MInt) (Biswas et al., 2015), Regularized Estimation of the Basis Covariance (REBACCA) (Ban et al., 2015) and Sparse Inverse Covariance Estimation for Ecological Association Inference (SPIEC-EASI) (Kurtz et al., 2015). Some of these tools have been evaluated using synthetic data (Weiss et al., 2016) or using data on plankton interactions reported from the literature (Lima-Mendez et al., 2015). Using these tools, microbial interactions in various environmental samples have been explored, such as marine water (Steele et al., 2011; Gilbert et al., 2012), fresh water (Eiler et al., 2012; Kara et al., 2013), the human microbiome (Qin et al., 2010; Arumugam et al., 2011; Duran-Pinedo et al., 2011; Faust et al., 2012; Greenblum et al., 2012; Endesfelder et al., 2014) and soils (Janssen, 2006; Zhou et al., 2011; Barberan et al., 2012; Lu et al., 2013; Hoppe et al., 2014). Networks inferred from each set of samples exhibited numerous links between numerous taxa. However, these links are often difficult to understand from biological and/or ecological perspectives. Previous studies have demonstrated that most networks inferred from environmental samples exhibit non-random patterns, but the patterns may or may not represent biological interactions. For example, two taxa may respond to soil temperature in a similar manner, and consequently result in similar abundance profiles, without being interaction partners. It is not clear whether the linked members are functionally associated taxa that work together to achieve a distinct function or an ecological process in the given environment. Indeed, interpreting such networks from biological/ecological perspectives is difficult, partially because of the complexity of microbial communities and partially because network inference tools cannot distinguish between true ecological interactions and other non-random processes (Faust and Raes, 2012). Consequently, abundance-based network inference tools have served more often as a powerful visual means to compare microbial communities across environmental gradients than as a means to appreciate microbial interactions associated with certain ecological processes.

To make better use of inference tools to appreciate microbial interactions in environmental samples, we propose the combinational use of different tools to capture tool-agreed networks. Different tools usually use different algorithms, methods to treat similarity scores and approaches to filter noises (false links). Some tools are designed to address particular computational questions. For example, eLSA (Ruan et al., 2006) and MInt (Biswas et al., 2015) are particularly suited to capture the time-lag relationships between two nodes (OTUs, taxa or species) in time-series data. MENA uses random matrix theory to determine thresholds on similarity scores (Zhou et al., 2010), whereas other tools determine thresholds by using permutation tests. CoNet is designed to use multiple similarity and distance measures (Pearson, Spearman, Bray Curtis, Kullback-Leibler and others) in combination to capture measure-agreed edges (links), and is based on the idea that different similarity measures would agree on true interactions and disagree on false interactions (Faust et al., 2012), whereas other tools use either the Pearson or Spearman correlation method. Owing to the different techniques employed, different tools often produce different networks, thus making it difficult for users to appreciate which result represents biological interactions in a given environment. Furthermore, it has been demonstrated that no single inference method performs optimally across all datasets, and the prediction accuracy can be significantly improved with multiple inference methods (Marbach et al., 2012). A recent evaluation also concluded that combining tools improves their performance (Weiss et al., 2016). Therefore, tool-agreed networks would be of higher value to capture ecologically meaningful interactions than tool-specific networks. Unfortunately, previous studies on soil samples have often relied on a single tool. If taxa in such tool-agreed networks provide sufficient phenotypic information, this information can be used to interpret the functionality of networks and relate it to a certain soil process. This approach allows exploration of whether a network represents a functional guild associated with a distinct soil process. In addition, highly linked (high degree) taxa in inferred networks have often been referred to as "keystone species" and interpreted as species playing important roles in communities. However, the degree-based keystone species are frequently claimed based on single inference method only. Whether such species are tool-dependent has not been investigated. This kind of information helps us further understand the degree-based keystone species in environmental samples.

In the present work, we used water-submerged paddy soils as a model ecosystem. Rice paddy is an ecosystem full of aerobic, microaerobic and anaerobic zones and interfaces in micro- and macro-scales. Such ecosystem supports the life of bacteria with diverse physiologies. It is also known that biogeochemical variables fluctuates over time (Ponnamperuma, 1972), which drive the fluctuations of bacterial abundances over time, and thus provide an opportunity for us to sample soils with sufficient abundance variations that are necessary for inferring reliable network. We used three tools, CoNet, MENA and eLSA, to infer networks on the basis of abundance profiles, partitioned them into modules (sub-networks, each comprising a group of nodes more densely connected to each other than to nodes outside the group), screened for those

in which ≥ 50% of nodes were assignable to genus/species only, captured tool-agreed modules and tested their robustness against the changes in tool parameters. Using the phenotypic information of the nodes, we found two modules representing functional guilds and each was associated with a distinct paddy soil process. We also tried to validate three interactions using co-culture methods and demonstrated that they were interacting partners in laboratory but the validated interaction types were different from those inferred.

2. Materials and methods

2.1. Soil sample description

The experiment is located in a well-controlled area of experimental farm in Wenhuilu campus, Yangzhou University (119.422N; 32.388E). The mean annual temperature ranges between 14.8 and 15.3 °C and annual precipitation between 961 and 1048 mm. The experiment included control and copper treatments designed in three replicates, but only the control treatment without copper amendment was used in the present study. The experiment was set up in 2006. The soil, loam in texture, is developed on alluvial deposit of Yangtze River. The top layer of soil from 0 to 45 cm was removed and piled aside. Cement tanks, designed with adjustable drainage system, were constructed. Each tank measures 3.8 m (length) × 1.3 m (width) × 0.45 m (depth), with a surface area 5.1 m^2 available for rice cultivation. The soil piled aside was air-dried, thoroughly mixed mechanically and re-packed back to tanks. This procedure ensured the homogeneity of soils in all replicate tanks. After a period of soil settlement, tanks were available for rice cultivation. Since 2007, rice has been cultivated one season per year, with a fertilizer amendment rate equivalent to 250 kg N ha-1, 75 kg P2O5 ha-1 and 75 kg K2O ha-1. Water submergence is generally maintained during rice seasons. All agronomic practices are done manually and applied to all tanks as uniformly as possible. The experimental design has been previously described by Zhao et al. (2012a, b). Soil samples were collected at weekly intervals, spanning an entire rice season from 10 June to 23 October 2014. Six soil cores (0–18 cm in depth) were collected from each tanks at each sampling time. Sub-samples were air-dried for soil properties measurements. Wet sub-samples were stored under −80 °C for later bacterial community analyses. Because soil properties measured (Fig. S1) were not significantly different between replicate tanks, samples from three replicate tanks were pooled into a composite before DNA extractions, which was mixed by hand. This procedure resulted in 17 samples each representing a time point. Insignificant differences between replicates were also evidenced by chemical speciation of copper in soils independently reported by Zhao et al. (2012b).

2.2. DNA extraction, amplicon sequencing and sequence processing

Bulk soil DNA was extracted using a Power Soil DNA Isolation Kit (Mobio Laboratories, Carlsbad, CA, USA) following the manufacturer's protocol. DNA quality was evaluated by 1% sodium boric acid agarose gel electrophoresis. DNA concentrations were determined using a NanoDrop 1000 spectrophotometer (Thermo Scientific, Waltham, MA, USA). The V4 hypervariable region of the 16S rRNA gene was amplified using the 563F (5′-AYTGGGYDTAAAGVG-3′) and 802R (5′-TACNVGGGTATCTAATCC-3′) primer pair (Cardenas et al., 2010) with appropriate barcodes. Each PCR mixture (20 μl) contained 5 × FastPfu buffer (4 μl), 2.5 mM dNTPs (2 μl), 5 μM of each primer with barcodes (0.4 μl), DNA (0.5 μl), FastPfu polymerase (0.4 μl), and ddH_2O (TransGen Biotech, Beijing, China). Amplification was performed using the following program: 95 °C for 2 min; 30 cycles of 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s; 72 °C for 5 min; hold at 10 °C. The amplicons were purified, and equal quantities of the purified PCR products were loaded for sequencing on a commercial MiSeq Illumina platform (Shanghai Majorbio Bio-Pharm Technology Co.,Ltd., Shanghai, China).

The sequence reads were processed using QIIME (v1.9.1) (Caporaso et al., 2010). The operational taxonomic unit (OTU) was assigned at a 97% identity level using the Usearch program (v7.0.1001) (Edgar, 2010). OTUs at 95%, 90% and 85% identity levels were similarly obtained but were used to study network reproducibility only. The OTUs were classified using the RDP database with the online version of the RDP classifier (Wang et al., 2007). After filtering, an average of 32 872 high-quality sequences per sample were obtained and used for subsequent analyses. The raw sequences have been submitted to the NCBI short read archive under the accession number SRP045267.

2.3. Generation of interaction networks

Three network inference tools, CoNet (v. 1.0.6 beta), MENA (web version) and eLSA (v. 7e618ed), were used to generate interaction networks. For CoNet (Faust et al., 2012), a read count matrix and a taxonomy matrix were loaded. The read count matrix was filtered by setting 15 as the minimum occurrence value across 17 samples (only two zero occurrences were allowed). Pair-wise associations among OTUs were calculated using the Pearson, Spearman, Kendall, Bray Curtis and Kullback-Leibler correlation methods simultaneously. The initial top and bottom edge numbers were set at 2000. For each edge and each measure of association, 1000 permutation scores and 1000 bootstrap scores were computed. For the permutation of the correlation measures, the vectors of taxon pairs were first shuffled and then renormalized before the scores were computed to mitigate compositional bias. An edge- and measure-specific p-value was obtained as the area under the bootstrap distribution limited by the mean of the permutation distribution (see reference Faust et al. (2012) for details regarding the p-value computation). The measure-specific p-values were then merged using Brown's method (Brown, 1975) and were corrected for multiple testing with the Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995). Edges supported by at least two correlation methods were retained. Edges with scores outside the limits of the 95% confidence interval defined by the bootstrap distribution or with adjusted p-values above 0.05 were discarded. A final network was restored from the permutation and bootstrap files.

For MENA (Deng et al., 2012), the read counts were filtered manually to exclude OTUs with > two “zero” occurrences across 17 samples and were uploaded to the pipeline. The majority was set to one, missing data were kept blank, read counts were converted by the logarithm, and the Pearson correlation coefficient was used as the similarity measure. For eLSA (Xia et al., 2013), the manually filtered data used for MENA were also used for eLSA analysis. The initial computation used the theoretical approximation as the p-value method and three as the time delay (using the command line “lsa_compute testna.txt result.txt -s 17 -p theo -x 1000 -d 3”, with a p-value cutoff at 0.05). The results were then filtered by keeping pair-wise correlations with absolute LA values ≥ 0.88. The general properties of the inferred networks were analysed using NetworkAnalyzer (Assenov et al., 2008). For simplicity, we used positive-edge-only sub-networks for subsequent analyses for all tool-derived networks. The positive-edge-only sub-networks were partitioned into modules using the OH-PIN algorithm (Wang et al., 2012) with the default settings (threshold = 2; overlapping score = 0.5). The minimal number of nodes was set at three. The networks were visualized in Cytoscape (v. 3.2.1) (Shannon et al., 2003).

2.4. Tool-agreed modules and their robustness test

In the present study, tool-agreed modules were defined as those that were reproducible across at least two tools and shared ≥50% of order-level nodes. To facilitate later functionality interpretation, we further limited the modules to those with ≥50% of nodes assignable to genus/species, because these nodes are phenotypically informative and some of them can be used for co-culture studies. For the robustness test, only CoNet was used. We adjusted CoNet's parameters (Table S1), constructed eight networks separately and computed their intersection networks using an in-house developed script. Edges that were consistently captured in the intersection networks across all taxonomic levels were recognized as the robust edges.

2.5. Co-culture experiments

To validate the direct interactions inferred, we performed co-culture experiments in the laboratory. Three edges, specifically between *Janthinobacterium lividum* and *Leuconostoc lactis*, between *J. lividum* and *Lactococcus piscium* and between *Leu. lactis* and *Lac. piscium*, were selected from a tool-agreed and robust module. *Leu. lactis* (CGMCC 1.2137) and *J. lividum* (CGMCC 2308) were purchased from the China General Microbiological Culture Collection Centre, Institute of Microbiology, Chinese Academy of Sciences, and *Lac. piscium* (ATCC 700018) was obtained from the American Type Culture Collection. We performed 16S rRNA gene sequencing in our laboratory to re-confirm that the species purchased were the same as those chosen on the basis of 16S rRNA gene sequences. Brain heart infusion medium (ATCC medium 44) (pH 7.0) was used throughout this experiment. The strains were pre-cultured overnight in liquid and their cell densities were determined using qPCR approach, based on which the inoculant volume of each strain was adjusted to achieve approximately 1:1 initial cell ratios in each set of co-culture experiment (ranging from 0.8:1 to 1.4:1 as practically measured). Strains were inoculated into 100 ml of fresh medium in a 150-ml flask either alone, mixed in pairs or mixed in three. The flask was then incubated at 25 °C without shaking, but was gently shaken by hand at the sampling time points. The growth curve was determined by quantitative PCR based on three replicate flasks. The primers for qPCR are given in Table S2.

3. Results

3.1. General properties of the networks

Using the abundance data, we constructed networks using CoNet, MENA and eLSA. The networks exhibited common scale-free and small-world properties (Fig. S2), and their power-law correlation coefficients (γ-values) and shortest path length values approximately matched the criteria proposed by (Barabasi and Oltvai, 2004) and were similar to the values of most ecological networks as summarized by Deng et al. (2012). Regarding the highly linked nodes, which are often referred to as “keystone species” or “hubs”, we found that different tools captured different sets of hubs (Fig. S2). Among the top 10 hubs at the OTU level, only four (OTU6892, OTU7287, OTU9695 and OTU9449) were conserved across the three tools. Four hubs (OTU2667, OTU7784, OTU8849 and OTU9522) were unique to CoNet, three (OTU292, OTU3243 and OTU3786) to MENA, and three (OTU8867, OTU4946 and OTU1551) to eLSA. Other hubs overlapped between two of the tools. The differences were observed at even the family or order levels. For example, an Alcaligenaceae-associated hub and a Candidate_division_WS3-associated hub were detected by CoNet only, and two Xanthobacteraceae-associated hubs were detected by eLSA only.

3.2. Tool-agreed and robust modules

The networks were then partitioned into modules. Taxonomic lineages of the nodes were examined. The modules derived from different tools, containing the shared majority (≥50%) of order-level nodes and consisting of nodes dominantly (≥50%) assignable to genus/species were defined as tool-agreed modules. Through this approach, we identified four tool-agreed modules, two of which were three-tool-agreed (Fig. 1) and the other two were two-tool-agreed (Fig. S3). The number of tool-agreed modules accounted for only a small fraction of the total modules inferred. Many modules were excluded because they did not meet these definitions. In the present study, totally 9806 OTUs were identified based on 97% similarity cut-off value, among which 2035 were assignable to genus and 198 to species. In the positive-only network, there were 413 nodes, 124 of which were genus-level taxa.

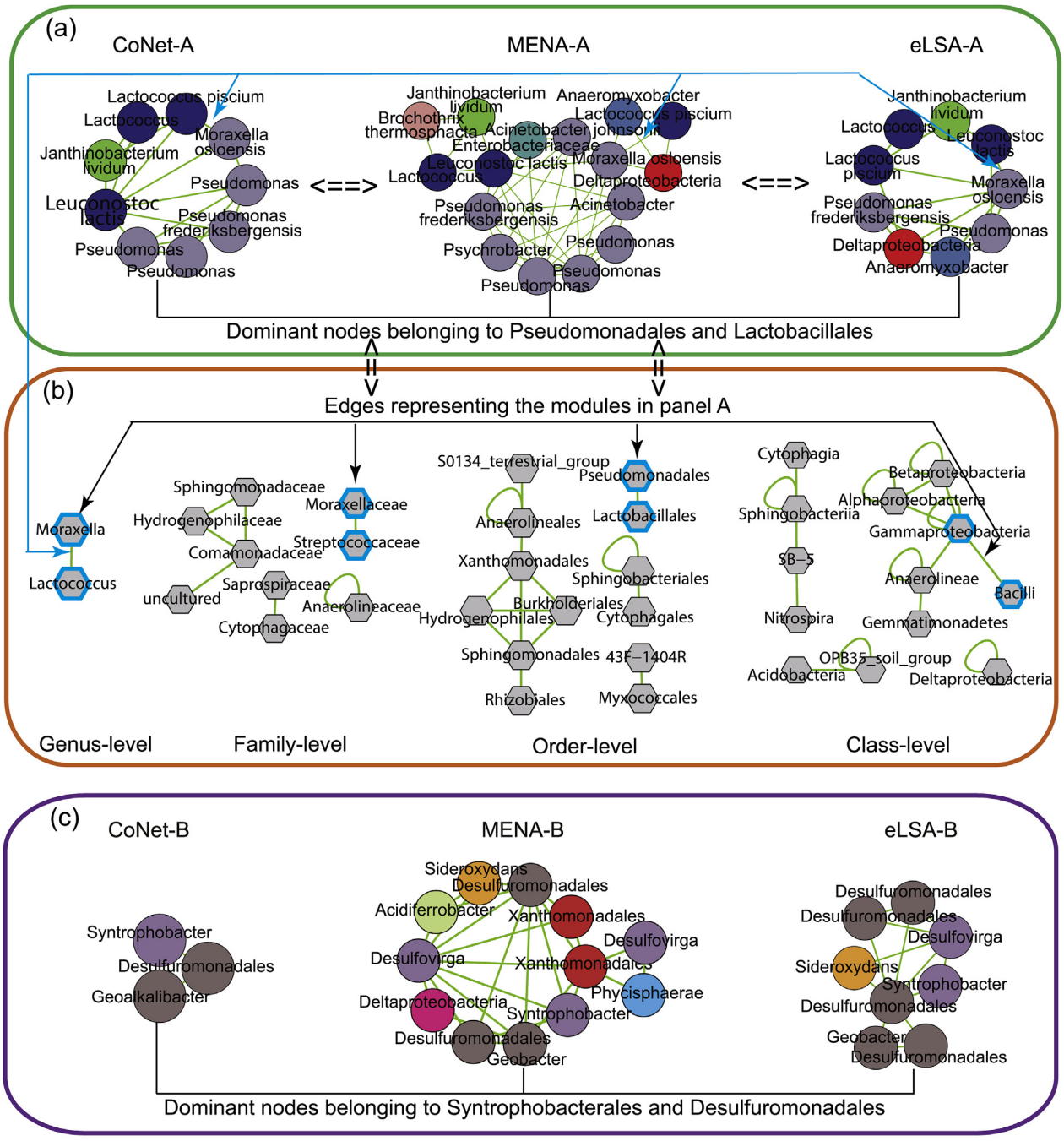
Next, we tested the robustness of the tool-agreed modules. The modules were considered robust if they were consistently present regardless of the change in the tool's parameters in a certain range. This time, we used only CoNet to test the robustness. Using a range of CoNet parameters (Table S1), we constructed eight networks, computed their intersection edges and examined the intersection edges across different taxonomic levels. In the genus-level intersection network, there was only one edge between *Moraxella* and *Lactococcus* (Fig. 1b). In the high-level intersection network, more edges were captured, and an edge taxonomically representing the genus-level edge was consistently found (labeled with a blue-coloured thick rim). For example, in the order-level intersection network, the edge between Pseudomonadales and Lactobacillales taxonomically represented the genus-level edge between *Moraxella* and *Lactococcus*. These two orders accounted for 68.8%–88.9% of the number of nodes in the modules of CoNet-A, MENA-A and eLSA-A (Fig. 1a). These data demonstrated that only one three-tool-agreed module (Fig. 1a) passed the robustness test. We did not find any other intersection edges that represented other tool-agreed modules (Fig. 1c and Fig. S3).

3.3. Interaction validations using co-culture methods

To validate the direct edges inferred, we selected three species for co-culture experiments: *Janthinobacterium lividum*, *Leuconostoc lactis* and *Lactococcus piscium*. These species were all from the tool-agreed and robust modules and the edges between them were all inferred to be positive (Fig. 1a). Co-culture was performed in two-species and three-species mixtures and results are summarized in the bottom-right panel of Fig. 2. In the two-species mixture, *Leu. lactis* grew at a higher rate in the presence of *J. lividum* than alone but *J. lividum* grew negligibly in the presence of *Leu. lactis*, showing an ambivalent interaction type in the laboratory (positive for one partner and negative for the other). In the three-species mixture, *J. lividum*, *Lac. piscium* and *Leu. lactis* all grew extremely poorer in the presence of the other two species than alone. The ambivalent interaction between *J. lividum* and *Leu. lactis* in the two-species mixture was not retained in the three-species mixture. These results showed that these species were interaction partners under the conditions used in the present study, but none of the interactions were validated as the same types as those inferred.

3.4. Interpretation of the potential ecological functions

After the above rigorous filtering procedures, we interpreted module functions and their associations with particular soil processes, on the basis of the phenotypic properties of the module members and on the general processes of paddy soils. Regarding the soil processes, it is well known that the shallow-water



Summary:

	CoNet	MENA	eLSA
Total No. of modules inferred	20	23	14
No. of modules with ⩾ 50% genus-/species-leveled nodes	7	5	5
No. of modules reproducible across three tools	2	2	2
No. of modules reproducible across two tools	1	1	2
No. of robust modules	1	1	1

Fig. 1. Three-tool-agreed modules and intersection networks. (a): Modules representing an association between aerobic respiration and fermentation. (b): Intersection networks capturing the edges (with a blue-coloured thick rim) that represent modules in panel (a) at various taxonomic levels. (c): Modules representing an association that completes metal/sulphur oxidation and reduction cycles. Nodes not assignable to species are labeled with the names of higher taxonomic ranks. Node colors in (a) and (c) represent orders. "⟨==⟩" indicates equivalence in node compositions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

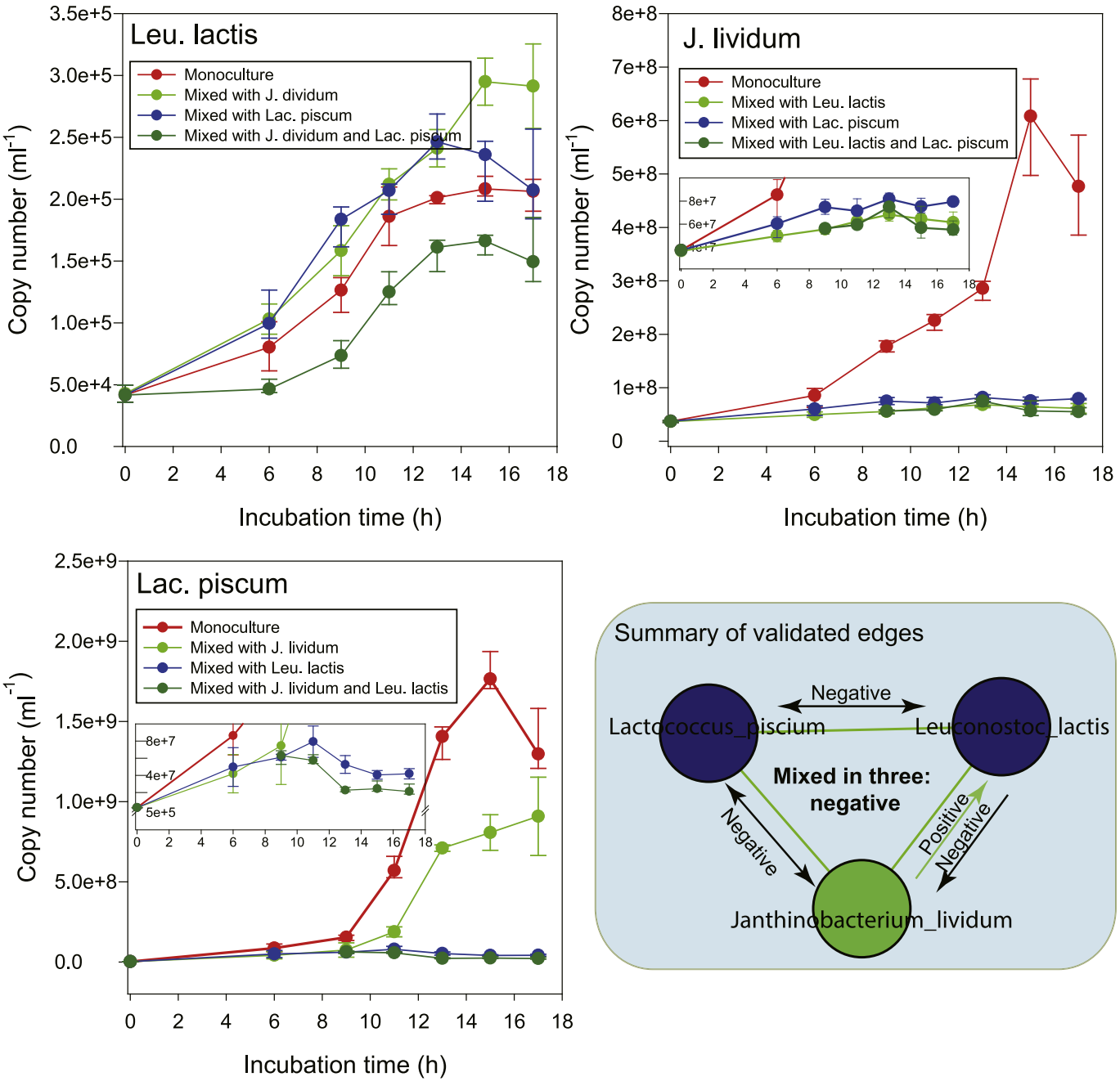


Fig. 2. Growth curve of each species in monoculture, two-species and three-species co-culture. The small incerted figures shows closeup details of y-axis. The bottom-right panel summarizes the results, where lines with arrows represent edges and their directions in the observed co-culture experiments. Green lines represent positive edges and black lines represent negetive edges. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

submerged paddy environment is full of aerobic, anoxic and anaerobic macro-/micro-sites, layers and interfaces (Ponnamperuma, 1972). Such a heterogeneous environment permits microbial aerobic respiration and fermentation to occur simultaneously and allows metal/sulphur oxidation/reduction cycles to complete. Regarding the taxa in specific modules, as described in the next two paragraphs, we found that their phenotypic properties were closely linked and were highly associated with distinct soil processes that requires joint efforts. Here, we considered only three-tool-agreed modules (Fig. 1a and c).

The modules CoNet-A, MENA-A and eLSA-A (Fig. 1a) exhibited a common feature: their members were either aerobes or facultative anaerobes with fermentative metabolism, despite their phylogenetic distances. CoNet-A consisted of nine nodes belonging to five defined genera, five of which represented cultivable species. The genera *Pseudomonas, Moraxella* (γ-Proteobacteria), and *Janthinobacterium* (δ-Proteobacteria) are common in their strictly aerobic and chemoorganotrophic metabolisms (Palleroni, 2005). The genera *Lactococcus* and *Leuconostoc* (Bacilli) are common in their facultative anaerobic and fermentative growth. The module eLSA-A was nearly identical to module CoNet-A in node composition, with only one additional node assignable to class δ-Proteobacteria. MENA-A had five more additional nodes belonging to four genera. These nodes were either aerobic or facultative as well. *Brochothrix*

(Bacilli) is a genus that is able to perform fermentative metabolisms (Sneath, 2009). *Acinetobacter* (Juni, 2005a) and *Psychrobacter* (γ-Proteobacteria) (Juni, 2005b) are described as aerobes. The node belonging to *Anaeromyxobacter* (δ-Proteobacteria) in MENA-A was the only exception that is a facultative anaerobe but cannot grow by fermentation (Sanford et al., 2002). Overall, this information suggested that these modules represented an association between aerobic respiration and fermentation. In this association, fermentation products, including a range of short-chain fatty acids (such as acetate, lactate, propionate, butyrate, etc.), may be used as substrates for aerobes and oxidized to CO_2 in the presence of O_2. Consequently, the quantity of fermentation products could be kept to a minimum, which is beneficial and necessary to sustain good rice growth (Oliveira de Sousa et al., 2002). Alternatively, fermentation products could be kept to a minimum by being converted to methane under strictly anaerobic environments (Glissmann and Conrad, 2000), but aerobic respiration is generally more energy efficient than methanogenesis (Thauer et al., 1977). Therefore, the association between aerobes or facultative anaerobes was meaningful from both microbiological and ecological perspectives. In short, the module apparently represented a functional guild that cooperated in the disposal of fermentation products.

The modules in Fig. 1c were three-tool-agreed but did not pass our robustness test. Nevertheless, their functions associated with soil processes can be appreciated. These modules also exhibited a common feature: their members were phenotypically relevant to metal/sulphur cycles. A node in CoNet-B belonged to the genus *Geoalkalibacter*, two species of which have been reported to use Fe(III), Mn(IV) or S^0 as electron acceptors (Zavarzina et al., 2006; Greene et al., 2009). Another node belonged to the genus *Syntrophobacter*, which uses sulphate as the electron acceptor when oxidizing propionate to acetate and CO_2 (Loka Bharathi, 2004; Muyzer and Stams, 2008). Modules in MENA-B and eLSA-B were more complicated in node compositions than CoNet-B, with several additional nodes that are also phenotypically relevant to metal/sulphur cycles. For instance, one additional node in both MENA-B and eLSA-B belonged to *Desulfovirga,* a genus that is able to disproportionate inorganic sulphur compounds to sulphate and sulphide (Kuever et al., 2005). Two nodes belonged to *Acidiferrobacter* and *Sideroxydans,* the type strains in both genera are able to oxidize Fe(II) (Emerson and Moyer, 1997; Hallberg et al., 2011) and the former is able to oxidize S^0, S^{2-} and $S_4O_6^{2-}$ in addition to Fe^{2+}. All of the above results indicated that the module represented a functional guild that cooperated to complete the oxidation and reduction cycles of $Mn^{4+}/Fe^{3+}/SO_4^{2-}/S^0$, which are essential processes in paddy soils. These cycles would not be completed if only oxidizers or reducers were present in the modules.

4. Discussion

4.1. Methodological considerations

A key feature of the present study is the use of three computational tools in combination to capture tool-agreed networks. The selected tools have different strengths and weaknesses. In the evaluation by Weiss et al. (2016), LSA was the tool that could most accurately detect ecological interactions in simulated data with not too many zeros. CoNet, together with another approach called MIC (Reshef et al., 2011), was the most robust tool, i.e. its results did not vary much when applied to repeatedly rarefied data sets. CoNet was also among the tools with the lowest false positive rate in simulated data. The choice of tools used in the present study is arbitrary, because we do not know which tool performs best on real-world sequencing data. A comparative evaluation of these tools on such data has not been carried out yet (but with simulated data only). At the present development stage, what we can do is to apply more than one network inference tool to analyse a real-world sequencing data set and to test a sub-set of the predictions experimentally. As for the number of tools to choose for combined uses, Marbach et al. (2012) have shown that performance can be significantly improved with as few as three methods. As for the prevalence cut-offs, different choice of this value certainly affects sensitivities. Higher the value is set, lower the number of taxa remains. However, there is currently no standard method on how to set the prevalence threshold. In fact, a zero in sequencing data is ambiguous. It may either represent a true absence or an unknown abundance below detection level. We selected 15 out of 17 such that similarity calculations based on too many zeros are avoided. The choice of prevalence value also depends on sampling strategies. Cross-sectional samples generally exhibit lower prevalence profiles than the time-series. Thus, a lower prevalence cut-off value to set up is encouraged for cross-sectional samples. While in this study we cannot solve such long-standing methodological challenges, we do show that tools with low accuracy on simulated data agreed on some edges and that edges selected based on tool agreement could be confirmed experimentally. These encouraging results suggest that the multi-tool approach may be a way to circumvent the low accuracy of individual tools.

4.2. "Noise" filtering

The combined use of these tools to detect tool-agreed networks can be viewed as a step of "noise" or error filtering that can not be filtered by individual tools. The robustness reflects how strongly a module is formed, which also can be viewed as a method of filtering. Another important filtering procedure is to screen for the modules with most ($\geq$50%) of nodes assignable to genus/species. This procedure is essential to interpret the modules' functions, which rely heavily on the phenotypic information provided by the nodes. Furthermore, this procedure is critical for the recognition of tool-agreed modules. Suppose that there are two modules, both containing nodes assignable to Proteobacteria (phylum) only. The probability of the modules to be tool-agreed is very low because the sequences of one module members assignable to that phylum may be tremendously different from those of other module members assignable to the same phylum. Indeed, many nodes/taxa in environmental communities are not assignable to genus/species. A module with such phenotypically non-informative nodes may have an ecological meaning but may not be appreciated and may not be useful for our purpose. This is why our approach captured only a few ecologically meaningful modules. To ultimately improve this situation, as many uncultured bacteria as possible should be cultured to update the databases used for classification and to provide more taxa with defined phenotypic information.

4.3. Experimental validation by the co-culture methods

Experimental validation is a necessary procedure to confirm true bacterial interactions in environmental samples. An interaction, no matter how well it is inferred by computational tools, cannot be confirmed as a true interaction unless it is experimentally validated. Unfortunately, previous studies have often inferred interactions without performing validation. In the present study, we validated three interactions using co-culture methods. The results showed that the species involved were interacting partners in the laboratory, but the validated interaction types were different from those inferred (Fig. 2), thus illustrating how difficult it is to investigate bacterial interactions in environmental samples. Currently, co-culture methods remain standard procedures to

validate interactions between microorganisms. These methods are advantageous in providing controlled conditions to test ecological concepts that cannot be easily tested in macro-ecological systems; however, these methods are disadvantageous in creating artificial conditions that do not represent natural environments and in creating an over-simplified community comprising only a few species removed from the whole complex community. Interactions occurring in such artificial environments are doomed to be different from those occurring in natural environments. Such differences can be observed in even the simplest co-culture system that we used. As demonstrated in Fig. 2, the ambivalent interaction between *J. lividum* and *Leu. lactis* shown in the two-species mixture was changed when a third party, *Lac. piscium*, was introduced. Similar results have previously been reported. For example, the addition of a non-susceptible decoy, *B. subtilis*, has been found to decrease the predation rate of *Bdellovibrio bacteriovorus* on its prey species *E. coli* S17-1 and to simultaneously result in partially compensatory enhancement of the availability of prey for predation (Hobley et al., 2006). The third-party effect has also been found in an inhibitory interaction between an antibiotic-producing and an antibiotic-sensitive species (Kelsic et al., 2015). In a four-strain community (Kato et al., 2008), the interspecies relationships have been found to be intertwined in a complex manner, and all of the interactions contribute to the structural stability of the community. Clearly, validating bacterial interactions in environmental samples is a challenge. To improve this situation, the development of novel co-culture methods able to measure interactions among a large number of cultivable and uncultivable species under natural or nearest to natural conditions is needed (Faust and Raes, 2012); however, such methods are not currently available.

4.4. "Keystone species"

Another interesting result revealed in the present study is that different computational tools inferred different sets of degree-based (number of links) "keystone species", despite similar topological properties of the networks (Fig. S2). These results were not revealed if only a single tool was used. Degree-based "keystone species" have often been cited as the species playing important roles in communities. However, our results underscore that caution must be taken in their interpretation. The concept of keystone species was originally proposed by Paine (1969), on the basis of species removal experiments but not the species' degree. Whether highly linked taxa are ecologically more relevant than less-connected taxa remains unclear (Faust and Raes, 2012). For example, in one study (Hartwell et al., 1999), the removal of low-degree species from food webs has been found to have a large effect on the community. Unfortunately, the relevance of this concept to natural microbial communities has never been tested through a species removal experiment.

Acknowledgements

The authors are grateful to Professor Yunxia Wang for discussions and background information regarding the field experiment and Mr. Xiaoyi Liang, who calculated the initial eLSA correlations. This work was supported by the National Key Basic Research Program of China (2015CB150503), the National Natural Science of China (41071117), the Natural Science Foundation of Jiangsu Province, China (BK20130438) and the Natural Science Fund for Colleges and Universities in Jiangsu Province (14KJ180025). K. Faust and J. Raes were supported by the Research Foundation Flanders, the Flemish agency for Innovation by Science and Technology (IWT), the EU-FP7 grant METACARDIS HEALTH-F4-2012-305312, and by KU Leuven and the Rega Institute.

Appendix A. Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.soilbio.2016.11.029.

References

Arumugam, M., Raes, J., Pelletier, E., Le Paslier, D., Yamada, T., Mende, D.R., Fernandes, G.R., Tap, J., Bruls, T., Batto, J.M., Bertalan, M., Borruel, N., Casellas, F., Fernandez, L., Gautier, L., Hansen, T., Hattori, M., Hayashi, T., Kleerebezem, M., Kurokawa, K., Leclerc, M., Levenez, F., Manichanh, C., Nielsen, H.B., Nielsen, T., Pons, N., Poulain, J., Qin, J., Sicheritz-Ponten, T., Tims, S., Torrents, D., Ugarte, E., Zoetendal, E.G., Wang, J., Guarner, F., Pedersen, O., de Vos, W.M., Brunak, S., Dore, J., Antolin, M., Artiguenave, F., Blottiere, H.M., Almeida, M., Brechot, C., Cara, C., Chervaux, C., Cultrone, A., Delorme, C., Denariaz, G., Dervyn, R., Foerstner, K.U., Friss, C., van de Guchte, M., Guedon, E., Haimet, F., Huber, W., van Hylckama-Vlieg, J., Jamet, A., Juste, C., Kaci, G., Knol, J., Lakhdari, O., Layec, S., Le Roux, K., Maguin, E., Merieux, A., Melo Minardi, R., M'Rini, C., Muller, J., Oozeer, R., Parkhill, J., Renault, P., Rescigno, M., Sanchez, N., Sunagawa, S., Torrejon, A., Turner, K., Vandemeulebrouck, G., Varela, E., Winogradsky, Y., Zeller, G., Weissenbach, J., Ehrlich, S.D., Bork, P., 2011. Enterotypes of the human gut microbiome. Nature 473, 174–180.

Assenov, Y., Ramirez, F., Schelhorn, S.E., Lengauer, T., Albrecht, M., 2008. Computing topological parameters of biological networks. Bioinformatics 24, 282–284.

Ban, Y., An, L., Jiang, H., 2015. Investigating microbial co-occurrence patterns based on metagenomic compositional data. Bioinformatics 31 (20), 3322–3329.

Barabasi, A.L., Oltvai, Z.N., 2004. Network biology: understanding the cell's functional organization. Nature Reviews Genetics 5, 101–113.

Barberan, A., Bates, S.T., Casamayor, E.O., Fierer, N., 2012. Using network analysis to explore co-occurrence patterns in soil microbial communities. ISME J 6, 343–351.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological) 57, 289–300.

Biswas, S., McDonald, M., Lundberg, D.S., Dangl, J.L., Jojic, V., 2015. Learning microbial interaction networks from metagenomic count data. In: Przytycka, M.T. (Ed.), Research in Computational Molecular Biology: 19th Annual International Conference, RECOMB 2015. Springer International Publishing, Cham, Warsaw, Poland, pp. 32–43. April 12-15, 2015, Proceedings.

Brown, M.B., 1975. A method for combining non-independent, one-sided tests of significance. Biometrics 31, 987–992.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I., Huttley, G.A., Kelley, S.T., Knights, D., Koenig, J.E., Ley, R.E., Lozupone, C.A., McDonald, D., Muegge, B.D., Pirrung, M., Reeder, J., Sevinsky, J.R., Turnbaugh, P.J., Walters, W.A., Widmann, J., Yatsunenko, T., Zaneveld, J., Knight, R., 2010. QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7, 335–336.

Cardenas, E., Wu, W.M., Leigh, M.B., Carley, J., Carroll, S., Gentry, T., Luo, J., Watson, D., Gu, B., Ginder-Vogel, M., Kitanidis, P.K., Jardine, P.M., Zhou, J., Criddle, C.S., Marsh, T.L., Tiedje, J.M., 2010. Significant association between sulfate-reducing bacteria and uranium-reducing microbial communities as revealed by a combined massively parallel sequencing-indicator species approach. Appl Environ Microbiol 76, 6778–6786.

Deng, Y., Jiang, Y.H., Yang, Y., He, Z., Luo, F., Zhou, J., 2012. Molecular ecological network analyses. BMC Bioinformatics 13, 113.

Duran-Pinedo, A.E., Paster, B., Teles, R., Frias-Lopez, J., 2011. Correlation network analysis applied to complex biofilm communities. PLoS One 6, e28438.

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26, 2460–2461.

Eiler, A., Heinrich, F., Bertilsson, S., 2012. Coherent dynamics and association networks among lake bacterioplankton taxa. ISME J 6, 330–342.

Emerson, D., Moyer, C., 1997. Isolation and characterization of novel iron-oxidizing bacteria that grow at circumneutral pH. Appl Environ Microbiol 63, 4784–4792.

Endesfelder, D., zu Castell, W., Ardissone, A., Davis-Richardson, A.G., Achenbach, P., Hagen, M., Pflueger, M., Gano, K.A., Fagen, J.R., Drew, J.C., Brown, C.T., Kolaczkowski, B., Atkinson, M., Schatz, D., Bonifacio, E., Triplett, E.W., Ziegler, A.G., 2014. Compromised gut microbiota networks in children with anti-islet cell autoimmunity. Diabetes 63, 2006–2014.

Fang, H., Huang, C., Zhao, H., Deng, M., 2015. CCLasso: correlation inference for compositional data through Lasso. Bioinformatics 31 (19), 3172–3180.

Faust, K., Raes, J., 2012. Microbial interactions: from networks to models. Nature Reviews Microbiology 10, 538–550.

Faust, K., Sathirapongsasuti, J.F., Izard, J., Segata, N., Gevers, D., Raes, J., Huttenhower, C., 2012. Microbial co-occurrence relationships in the human microbiome. PLoS Comput Biol 8, e1002606.

Fisher, C.K., Mehta, P., 2014. Identifying keystone species in the human gut microbiome from metagenomic timeseries using sparse linear regression. PLoS One 9, e102451.

Friedman, J., Alm, E.J., 2012. Inferring correlation networks from genomic survey data. PLoS Comput Biol 8, e1002687.

Gilbert, J.A., Steele, J.A., Caporaso, J.G., Steinbruck, L., Reeder, J., Temperton, B., Huse, S., McHardy, A.C., Knight, R., Joint, I., Somerfield, P., Fuhrman, J.A., Field, D.,

2012. Defining seasonal marine microbial community dynamics. ISME J 6, 298–308.
Glissmann, K., Conrad, R., 2000. Fermentation pattern of methanogenic degradation of rice straw in anoxic paddy soil. FEMS Microbiology Ecology 31, 117–126.
Greenblum, S., Turnbaugh, P.J., Borenstein, E., 2012. Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. Proceedings of the National Academy of Sciences of the United States of America 109, 594–599.
Greene, A.C., Patel, B.K.C., Yacob, S., 2009. Geoalkalibacter subterraneus sp. nov., an anaerobic Fe(III)- and Mn(IV)-reducing bacterium from a petroleum reservoir, and emended descriptions of the family Desulfuromonadaceae and the genus Geoalkalibacter. International Journal of Systematic and Evolutionary Microbiology 59, 781–785.
Hallberg, K.B., Hedrich, S., Johnson, D.B., 2011. Acidiferrobacter thiooxydans, gen. nov. sp. nov.; an acidophilic, thermo-tolerant, facultatively anaerobic iron- and sulfur-oxidizer of the family Ectothiorhodospiraceae. Extremophiles 15, 271–279.
Hartwell, L.H., Hopfield, J.J., Leibler, S., Murray, A.W., 1999. From molecular to modular cell biology. Nature 402, C47–C52.
Hobley, L., King, J.R., Sockett, R.E., 2006. Bdellovibrio predation in the presence of decoys: three-way bacterial interactions revealed by mathematical and experimental analyses. Appl Environ Microbiol 72, 6757–6765.
Hoppe, B., Kahl, T., Karasch, P., Wubet, T., Bauhus, J., Buscot, F., Krüger, D., 2014. Network analysis reveals ecological links between N-fixing bacteria and wood-decaying fungi. PLoS One 9, e88141.
Janssen, P.H., 2006. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. Appl Environ Microbiol 72, 1719–1728.
Juni, E., 2005a. Genus acinetobacter. In: Brenner, D.J., Krieg, N.R., Staley, J.T., Garrity, G.M. (Eds.), Bergey'S Manual of Systematic Bacteriology, 2, Part B. Springer-Verlag, New York, N.Y, pp. 425–437.
Juni, E., 2005b. Genus psychrobacter. In: Brenner, D.J., Krieg, N.R., Staley, J.T., Garrity, G.M. (Eds.), Bergey'S Manual of Systematic Bacteriology, 2, Part B. Springer-Verlag, New York, N.Y, pp. 437–441.
Kara, E.L., Hanson, P.C., Hu, Y.H., Winslow, L., McMahon, K.D., 2013. A decade of seasonal dynamics and co-occurrences within freshwater bacterioplankton communities from eutrophic Lake Mendota, WI, USA. ISME J 7, 680–684.
Kato, S., Haruta, S., Cui, Z.J., Ishii, M., Igarashi, Y., 2008. Network relationships of bacteria in a stable mixed culture. Microbial Ecology 56, 403–411.
Kelsic, E.D., Zhao, J., Vetsigian, K., Kishony, R., 2015. Counteraction of antibiotic production and degradation stabilizes microbial communities. Nature 521, 516–519.
Kuever, J., Rainey, F.A., Widdel, F., 2005. Genus desulfocapsa. In: Brenner, D.J., Krieg, N.R., Staley, J.T., Garrity, G.M. (Eds.), Bergey'S Manual of Systematic Bacteriology, vol. 2, Part C. Springer-Verlag, New York, N.Y, pp. 992–994.
Kurtz, Z.D., Muller, C.L., Miraldi, E.R., Littman, D.R., Blaser, M.J., Bonneau, R.A., 2015. Sparse and compositionally robust inference of microbial ecological networks. PLoS Computational Biology 11, e1004226.
Lima-Mendez, G., Faust, K., Henry, N., Decelle, J., Colin, S., Carcillo, F., Chaffron, S., Ignacio-Espinosa, J.C., Roux, S., Vincent, F., Bittner, L., Darzi, Y., Wang, J., Audic, S., Berline, L., Bontempi, G., Cabello, A.M., Coppola, L., Cornejo-Castillo, F.M., d'Ovidio, F., De Meester, L., Ferrera, I., Garet-Delmas, M.-J., Guidi, L., Lara, E., Pesant, S., Royo-Llonch, M., Salazar, G., Sánchez, P., Sebastian, M., Souffreau, C., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Gorsky, G., Not, F., Ogata, H., Speich, S., Stemmann, L., Weissenbach, J., Wincker, P., Acinas, S.G., Sunagawa, S., Bork, P., Sullivan, M.B., Karsenti, E., Bowler, C., de Vargas, C., Raes, J., 2015. Determinants of community structure in the global plankton interactome. Science 348 (6237), 1262073.
Loka Bharathi, P.A., 2004. Synergy in sulfur cycle: the biogeochemical significance of sulfate reducing bacteria in syntrophic associations. In: Ramaiah, N. (Ed.), Marine Microbiology: Facets & Opportunities. National Institute of Oceanography, Goa, pp. 27–35.
Lu, L., Yin, S., Liu, X., Zhang, W., Gu, T., Shen, Q., Qiu, H., 2013. Fungal networks in yield-invigorating and -debilitating soils induced by prolonged potato monoculture. Soil Biology and Biochemistry 65, 186–194.
Marbach, D., Costello, J.C., Küffner, R., Vega, N.M., Prill, R.J., Camacho, D.M., Allison, K.R., The DREAM5 Consortium, Kellis, M., Collins, J.J., Stolovitzky, G., 2012. Wisdom of crowds for robust gene network inference. Nature Methods 9, 796–804.
Muyzer, G., Stams, A.J.M., 2008. The ecology and biotechnology of sulphate-reducing bacteria. Nature Reviews Microbiology 6, 441–454.
Oliveira de Sousa, R., do Carmo Ruaro Peralba, M., Meurer, E.J., 2002. Short chain organic acid dynamics in solution of flooded soil treated with ryegrass residues. Communications in Soil Science and Plant Analysis 33, 779–787.
Paine, R.T., 1969. A note on trophic complexity and community stability. American Naturalist 103.
Palleroni, N.J., 2005. Genus Pseudomonas. In: Brenner, D.J., Krieg, N.R., Staley, J.T., Garrity, G.M. (Eds.), Bergey'S Manual of Systematic Bacteriology, vol. 2, Part B. Springer-Verlag, New York, N.Y, pp. 323–379.
Ponnamperuma, F.N., 1972. The chemistry of submerged soils. Advances in Agronomy 24, 29–96.
Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D.R., Li, J., Xu, J., Li, S., Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., Xie, Y., Tap, J., Lepage, P., Bertalan, M., Batto, J.M., Hansen, T., Le Paslier, D., Linneberg, A., Nielsen, H.B., Pelletier, E., Renault, P., Sicheritz-Ponten, T., Turner, K., Zhu, H., Yu, C., Jian, M., Zhou, Y., Li, Y., Zhang, X., Qin, N., Yang, H., Wang, J., Brunak, S., Dore, J., Guarner, F., Kristiansen, K., Pedersen, O., Parkhill, J., Weissenbach, J., Bork, P., Ehrlich, S.D., 2010. A human gut microbial gene catalogue established by metagenomic sequencing. Nature 464, 59–65.
Reshef, D.N., Reshef, Y.A., Finucane, H.K., Grossman, S.R., McVean, G., Turnbaugh, P.J., Lander, E.S., Mitzenmacher, M., Sabeti, P.C., 2011. Detecting novel associations in large data sets. Science 334, 1518–1524.
Ruan, Q., Dutta, D., Schwalbach, M., Steele, J., Fuhrman, J., Sun, F., 2006. Local similarity analysis reveals unique associations among marine bacterioplankton species and environmental factors. Bioinformatics 22, 2532–2538.
Sanford, R.A., Cole, J.R., Tiedje, J.M., 2002. Characterization and description of Anaeromyxobacter dehalogenans gen. nov., sp. nov., an aryl-halorespiring facultative anaerobic myxobacterium. Appl Environ Microbiol 68, 893–900.
Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Research 13, 2498–2504.
Sneath, P.H.A., 2009. Genus brochothrix. In: De Vos, P., Garrity, G.M., Jones, D., Krieg, N.R., Ludwig, W., Rainey, F.A., Schleifer, K.-H., Whitman, W.B. (Eds.), Bergey'S Manual of Systematic Bacteriology, vol. 3. Springer-Verlag, New York, N.Y, pp. 257–268.
Steele, J.A., Countway, P.D., Xia, L., Vigil, P.D., Beman, J.M., Kim, D.Y., Chow, C.E., Sachdeva, R., Jones, A.C., Schwalbach, M.S., Rose, J.M., Hewson, I., Patel, A., Sun, F., Caron, D.A., Fuhrman, J.A., 2011. Marine bacterial, archaeal and protistan association networks reveal ecological linkages. ISME J 5, 1414–1425.
Thauer, R.K., Jungermann, K., Decker, K., 1977. Energy conservation in chemotrophic anaerobic bacteria. Bacteriological Reviews 41, 100–180.
Wang, J., Ren, J., Li, M., Wu, F.-X., 2012. Identification of hierarchical and overlapping functional modules in PPI networks. NanoBioscience, IEEE Transactions on 11, 386–393.
Wang, Q., Garrity, G.M., Tiedje, J.M., Cole, J.R., 2007. Naïve bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology 73, 5261–5267.
Weiss, S., Van Treuren, W., Lozupone, C., Faust, K., Friedman, J., Deng, Y., Xia, L.C., Xu, Z.Z., Ursell, L., Alm, E.J., Birmingham, A., Cram, J.A., Fuhrman, J.A., Raes, J., Sun, F., Zhou, J., Knight, R., 2016. Correlation detection strategies in microbial data sets vary widely in sensitivity and precision. ISME Journal 10, 1669–1681.
Xia, L.C., Ai, D., Cram, J., Fuhrman, J.A., Sun, F., 2013. Efficient statistical significance approximation for local similarity analysis of high-throughput time series data. Bioinformatics 29, 230–237.
Zavarzina, D.G., Kolganova, T.V., Bulygina, E.S., Kostrikina, N.A., Turova, T.P., Zavarzin, G.A., 2006. *Geoalkalibacter ferrihydriticus* gen. nov., sp. nov., the first alkaliphilic representative of the family *Geobacteraceae*, isolated from a soda lake. Mikrobiologiia 75, 775–785.
Zhao, J.N., Wang, Y.X., Shen, C.X., Yang, L.X., Wang, Y.L., 2012a. Effect of soil copper contamination on rice yield formation: a 5-year located experiment. Journal of Agro-Environment Science 31, 2073–2081.
Zhao, J.N., Wang, Y.X., Shen, C.X., Yang, L.X., Wang, Y.L., 2012b. Dynamic variations of distribution forms of exogenous copper in soil: a 5-years located experiment. Journal of Agro-Environment Science 31, 926–933.
Zhou, J., Deng, Y., Luo, F., He, Z., Tu, Q., Zhi, X., 2010. Functional molecular ecological networks. MBio 1, e00169, 10.
Zhou, J., Deng, Y., Luo, F., He, Z., Yang, Y., 2011. Phylogenetic molecular ecological network of soil microbial communities in response to elevated CO2. MBio 2, e00122, 11.